

Cite No 1.

China Patent Publication No.CN1378750A

[English abstract]

Apparatus, and associated method, for forming a compressed motion vector field utilizing predictive motion coding

Abstract

Apparatus, and an associated method, motion compensates coding of video sequences. Motion compensated prediction is utilized in the representation of motion vector fields. Reduced numbers of bits are required to represent the motion vector field while maintaining a low prediction error, thereby facilitating improved communication of, and recreation of, video frames forming a video sequence.

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.⁷

H04N 7/26

[12] 发明专利申请公开说明书

[21] 申请号 00814037.5

[43] 公开日 2002 年 11 月 6 日

[11] 公开号 CN 1378750A

[22] 申请日 2000.8.10 [21] 申请号 00814037.5

[30] 优先权

[32] 1999.8.11 [33] US [31] 09/371,641

[86] 国际申请 PCT/US00/21823 2000.8.10

[87] 国际公布 WO01/11891 英 2001.2.15

[85] 进入国家阶段日期 2002.4.8

[71] 申请人 诺基亚有限公司

地址 芬兰埃斯波

[72] 发明人 M·卡策维茨 J·莱尼曼

B·-P·多布林

[74] 专利代理机构 中国专利代理(香港)有限公司

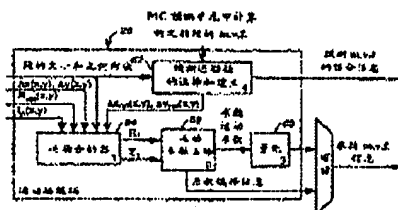
代理人 杨凯 傅康

权利要求书 14 页 说明书 30 页 附图 3 页

[54] 发明名称 用于压缩运动矢量场的装置和方法

[57] 摘要

视频序列的运动补偿编码的装置和相关方法。运动补偿预测被用于运动矢量场的表示。需要减少数量的比特来表示运动矢量场,同时保持低预测误差,由此便于改进形成视频序列的视频帧的传送和重建。



知识产权出版社出版

00814037.5

权 利 要 求 书

第1/14页

1. 在一种操作视频序列的方法中,所述视频序列包括至少当前视频帧和参考视频帧,所述当前视频帧包括至少一个第一相邻段和第二相邻段,用于所述当前帧的运动补偿预测的方法的改进包括:

检索之前储存的第一运动场模型,所述第一运动场模型是第一运动矢量场的模型,它关于所述参考视频帧中的象素来描述所述第一相邻段中象素的位移;

- 10 确定第二运动矢量场,它关于所述参考视频帧中的象素来描述所述当前视频帧的所述第二相邻段中的象素的位移;

利用运动模型来模拟所述第二运动矢量场,以形成第二运动场模型;

根据所述第一运动场模型来近似所述第二运动场模型,以形成预测场模型;

- 15 将所述第二运动场模型与所述预测场模型进行比较,并形成求精场(refinement field)模型,所述求精场模型表示所述第二运动场模型和所述预测场模型之间的差异;

通过对所述预测场模型和所述求精场模型求和来建立所述第二运动场模型的其它模型表示;

- 20 计算第一成本函数(cost function),其中所述第一成本函数包括遭受的第一图像失真的量度(measure)和在使用所述第二运动场模型时所需的第一数据量的量度;

- 25 计算第二成本函数,其中所述第二成本函数包括遭受的第二图像失真的量度和在使用所述第二运动场的所述其它模型表示时所需的第二数据量的量度;

将所述第一和第二成本函数进行比较,并确定所述第一和第二成本函数中哪一个具有较小绝对值;

在所述第二运动矢量场模型和所述第二运动矢量场的所述其它模

00814037.5

权 利 要 求 书 第2/14页

型表示之间选择与所述较小绝对值相关联的供选择的一个，以指示被选运动场模型，并储存所述被选运动场模型。

2. 如权利要求 1 所述的方法，其特征在于还包括：

对有关所述被选运动场模型的信息进行编码。

5 3. 如权利要求 2 所述的方法，其特征在于还包括：

向解码器发送所述编码信息，用于解码。

4. 如权利要求 2 所述的方法，其特征在于还包括：

将所述编码信息储存在存储装置中。

5. 如权利要求 1 所述的方法，其特征在于所述第一运动场模型、
10 所述第二运动场模型、所述第二运动场模型、所述预测场模型和所述求精场模型中的各个都被构成为运动场基本函数的和，每一个所述运动场基本函数都与某个运动系数相乘。

6. 如权利要求 5 所述的方法，其特征在于所述运动场基本函数是正交函数。

15 7. 如权利要求 6 所述的方法，其特征在于所述第一运动场模型、所述第二运动场模型、所述预测场模型和所述求精场模型中的各个都是仿射运动场模型。

8. 如权利要求 1 所述的方法，其特征在于所述至少一个第一相邻段和所述第二相邻段是四边形的。

20 9. 如权利要求 1 所述的方法，其特征在于还包括：

将所述至少一个第一相邻段分成多个子段，并利用至少一个所述子段的运动场模型以形成所述预测场模型。

10. 如权利要求 1 所述的方法，其特征在于所述预测场模型是通过对所选至少一个相邻段的运动场模型进行投影来形成的。

25 11. 如权利要求 1 所述的方法，其特征在于所述预测场模型是通过对一个以上的第一相邻段确定的所述第二运动矢量场的近似求平均来形成的。

12. 如权利要求 1 所述的方法，其特征在于所述预测场模型是通

00814037.5

权 利 要 求 书 第3/14页

过对从一个以上的第一相邻段确定的所述第二场模型的近似求平均来形成的。

13. 如权利要求 1 所述的方法，其特征在于计算所述第一成本函数的所述步骤是利用拉格朗日准则 (Lagrangian criterion) 来执行的。

5 14. 如权利要求 13 所述的方法，其特征在于所述拉格朗日准则具有 $L=D+\lambda \cdot B$ 的形式，其中 D 是在对给定运动系数组进行编码时遭受的失真，B 是表示所述运动系数所需的比特数， λ 是乘法拉格朗日参数 (multiplying Lagrangian parameter)。

15 15. 如权利要求 1 所述的方法，其特征在于利用共同的基本函数组来表示所述预测运动场和所述求精运动场。

16. 如权利要求 1 所述的方法，其特征在于还包括：

确定第一阈值；

识别在所述求精场模型的所有运动系数中具有最小值的所述求精场模型的运动的系数；

15 确定通过把所述最小运动系数设置为零所遭受的第三成本函数；

在所述第三图像失真没有超过所述第一阈值的情况下，通过把所述最小值的运动系数设置为零来形成所述求精场的近似；

17. 如权利要求 1 所述的方法，其特征在于如果所述被选运动场模型是所述第二运动场模型，则所述方法还包括：

20 把所述预测场模型的所有运动系数设置为零；

把所述求精场模型的所有运动系数设置为等于所述第二运动场模型的所述运动系数。

18. 如权利要求 17 所述的方法，其特征在于所述对信息编码以依赖于所述被选场模型的方式发生。

25 19. 如权利要求 18 所述的方法，其特征在于如果所述被选场模型是所述第二运动场模型，则所述对信息编码包括对所述求精场模型进行编码的步骤。

20. 如权利要求 18 所述的方法，其特征在于如果所述被选场模型

00814037.5

权 利 要 求 书 第4/14页

是所述其它模型表示,则所述对信息编码包括以下步骤:

对所述预测场模型进行编码;

对所述求精场模型进行编码。

21. 如权利要求 20 所述的方法,其特征在于所述对所述求精场模型进行编码包括以下步骤:

通过对第一和第二值中供选择的一个设置运动系数指示符,指示所述编码信息包括所述求精场模型的所述运动系数;

通过设置运动系数模式指示符,指示所述运动系数中那些具有非零值;

- 10 对所述非零运动系数值进行编码。

22. 如权利要求 21 所述的方法,其特征在于通过指示幅值和符号来对各个所述非零运动系数值进行编码。

23. 如权利要求 20 所述的方法,其特征在于对所述预测场模型进行编码包括以下步骤:

- 15 通过对第一和第二值中供选择的一个设置运动系数指示符,指示所述编码信息不包括运动系数值;

通过设置方向鉴别指示符,指示关于所述至少一个第一相邻段的所述第二相邻段的建立所述其它模型表示的方向。

- 20 24. 如权利要求 23 所述的方法,其特征在于对所述预测场模型进行编码还包括以下步骤:

通过设置子段鉴别指示符,指示建立所述其它模型表示的所述至少一个第一相邻段的子段。

- 25 25. 在一种操作视频序列的方法中,所述视频序列包括至少当前视频帧和参考视频帧,所述当前视频帧包括至少一个第一相邻段和第二相邻段,用于所述当前视频帧的运动补偿预测的方法的改进包括:

检索至少一个之前储存的第一运动场模型,所述至少一个第一运动场模型是第一运动矢量场的模型,它关于所述参考视频帧中的象素来描述所述至少一个第一相邻段中象素的位移;

00814037.5

权 利 要 求 书 第5/14页

确定第二运动矢量场，它关于所述参考视频帧中的象素来描述所述当前视频帧的所述第二相邻段中的象素的位移；

利用运动模型来模拟所述第二运动矢量场，以形成第二运动场模型；

- 5 根据所述至少一个第一运动场模型来近似所述第二运动场模型，以形成预测场模型；

26. 在一种用于对包括至少一个当前视频帧的视频序列进行操作的视频设备中，所述当前视频帧具有至少一个第一相邻段和第二相邻段，一种用于形成所述第二相邻段的运动矢量场的近似的装置的改进，所述装置包括：

- 10 进，所述装置包括：

运动矢量场建立器，它被连接以接收表示第一仿射运动模型指示并接收所述第二相邻段的指示，所述第一仿射运动模型形成表示所述第一相邻段的第一运动矢量场的近似，所述运动矢量场建立器用于对表示所述第一仿射运动模型的指示进行响应而形成第二仿射运动模型，所述第二仿射运动模型形成所述第二相邻段的运动矢量场的近似。

- 20 27. 如权利要求 26 所述的装置，其特征在于所述视频序列还包括参考视频帧，所述运动矢量场建立器还被连接以接收所述参考视频帧的指示，所述第二仿射运动模型是对表示所述第一仿射模型和所述参考视频帧的被选部分的指示中供选择的一个进行响应。

28. 如权利要求 27 所述的装置，其特征在于所述运动矢量场建立器还计算第二运动矢量场，计算的所述第二运动矢量场响应所述参考视频帧的所述被选部分。

- 25 29. 如权利要求 28 所述的装置，其特征在于所述运动矢量场建立器还确定所述第二运动矢量场和所述第二仿射运动模型之间的差异，其差异构成求精场模型。

30. 如权利要求 29 所述的装置，其特征在于所述运动矢量场建立器还建立所述第二运动矢量场的其它表示模型，所述第二运动矢量场

00814037.5

权 利 要 求 15 第6/14页

的所述其它表示模型是由所述求精场模型和所述第二仿射运动模型的组合构成的。

31. 如权利要求 30 所述的装置, 其特征在于所述运动矢量场建立器还确定成本函数, 所述成本函数至少部分地是与所述第二运动矢量场和所述第二仿射运动模型中至少一个被选的有关的数据要求和图像失真的表示。

32. 如权利要求 31 所述的装置, 其特征在于所述运动矢量场建立器还利用所述第二运动矢量场和所述第二仿射运动模型中被选的一个, 所作选择响应于所述成本函数。

33. 如权利要求 27 所述的装置, 其特征在于所述运动矢量场建立器还选择表示所述第一仿射模型和所述参考视频帧的所述被选部分的指示中供选择的一个, 所述第二仿射运动模型是对此进行响应而形成的。

34. 如权利要求 26 所述的装置, 其特征在于所述第一仿射运动模型具有与其相关联的第一仿射运动系数, 所述运动矢量场建立器还对所述第一仿射运动模型的值进行投影以形成所述第二仿射运动模型。

35. 如权利要求 26 所述的装置, 其特征在于所述当前视频帧还具有第三相邻段, 所述第三相邻段与所述第一相邻段和所述第二相邻段相邻, 所述运动矢量场建立器还被连接以接收所述第二相邻段的指示, 所述运动矢量场建立器还用于响应所述第一仿射运动模型和所述第二仿射运动模型中被选的供选择的一个, 而形成第三仿射运动模型。

36. 如权利要求 35 所述的装置, 其特征在于所述运动矢量场建立器还选择所述第一仿射运动模型和所述第二仿射运动模型中所述供选择的一个, 所述第三仿射运动模型是对此进行响应而形成的。

37. 如权利要求 26 所述的装置, 其特征在于所述视频设备构成视频序列发生器, 所述视频序列发生器具有编码器, 并且所述运动矢量场建立器构成所述编码器的一部分。

00814037.5

权 利 要 求 书 第7/14页

38. 如权利要求 26 所述的装置, 其特征在于所述视频设备构成视频序列接收机, 所述视频序列接收机具有解码器, 并且所述运动场建立器构成所述解码器的一部分。

39. 在一种操作视频序列的方法中, 所述视频序列包括至少一个
5 当前视频帧, 所述当前视频帧具有至少一个第一相邻段和第二相邻段, 一种用于形成所述第二相邻段的运动矢量场的近似的方法的改进, 所述方法包括:

形成表示所述第一相邻段的第一运动矢量场;

用第一仿射运动模型来作所述第一运动矢量场的模型; 以及

10 对所述建模操作期间建模的所述第一运动矢量场进行响应, 形成第二仿射运动矢量模型, 所述第二仿射运动模型形成所述第二相邻段的运动矢量场的近似。

40. 如权利要求 39 所述的方法, 其特征在于所述当前视频帧还包括第三相邻段, 所述第三相邻段与所述第一相邻段和所述第二相邻段
15 相邻, 所述方法还用于形成所述第三相邻段的运动矢量场的近似, 所述方法包括以下操作:

响应于所述第一仿射运动模型和第二仿射运动模型中供选择的一个而形成第三仿射运动模型。

41. 如权利要求 39 所述的方法, 其特征在于所述视频序列还包括
20 参考视频帧, 并且在形成所述第二仿射运动模型的所述操作期间形成的所述第二仿射运动矢量场是对所述第一运动矢量场和所述参考帧的一部分中供选择的一个作出响应。

42. 如权利要求 41 所述的方法, 其特征在于还包括以下的附加操作: 选择所述第一运动矢量场和所述参考帧中的一个, 所述第二仿射
25 运动模型是对此进行响应而形成的。

43. 在一种用于对包括至少一个当前视频帧和参考视频帧的视频序列进行操作的视频设备中, 所述当前视频帧具有至少一个第一相邻段和第二相邻段, 一种用于形成运动矢量场的近似的装置的改进, 所

00814037.5

权 利 要 求 书 第8/14页

述装置包括:

运动矢量场建立器, 它被连接以接收表示所述第一相邻段和所述第二相邻段中被选的一个的指示及表示所述参考视频帧的部分的指示, 所述运动矢量场建立器用于确定所述第一和第二相邻段中所述被选的一个和所述参考视频帧的被选部分之间的映射, 并用于用仿射运动模型来近似所述映射, 所述仿射运动模型形成所述运动矢量场的近似。

44. 如权利要求 43 所述的装置, 其特征在于所述第一相邻段和所述第二相邻段中所述被选的一个包括所述第一相邻段, 并且形成所述运动矢量场的所述近似的所述仿射运动模型包括第一仿射运动模型, 所述第一仿射运动模型表示所述第一相邻段。

45. 如权利要求 43 所述的装置, 其特征在于所述第一相邻段和所述第二相邻段中所述被选的一个还包括所述第二相邻段, 所述运动矢量场建立器还用于确定所述第二相邻段和所述参考视频帧的所述被选部分与所述第一相邻段中供选择的一个之间的映射, 并且形成所述运动矢量场的近似的所述仿射运动模型还包括第二仿射运动模型, 所述第二仿射运动模型表示所述第二相邻段。

46. 如权利要求 43 所述的装置, 其特征在于所述视频设备包括视频序列发生器, 所述视频序列发生器具有编码器, 并且所述运动矢量场建立器包括编码器的一部分。

47. 如权利要求 43 所述的装置, 其特征在于所述视频装置包括视频序列接收机, 所述视频序列接收机具有解码器, 并且所述运动矢量场建立器包括所述解码器的一部分。

48. 在一种对视频序列进行解码的方法中, 所述视频序列包括至少一个当前视频帧和参考帧, 所述当前帧包括至少一个第一相邻段和第二相邻段, 一种用于对所述当前视频帧进行解码的方法的改进包括以下步骤:

接收信息类型的指示;

00814037.5

权 利 要 求 书 第9/14页

接收所述第二相邻段的段重建信息;

响应所述指示而选择段重建模式;

按照所述被选段重建模式来重建所述第二相邻段。

49. 如权利要求 48 所述的方法, 其特征在于所述被选段重建模式
5 是一组段重建模式中的一种, 所述段重建模式组包括:

第一段重建模式, 其中所述段重建信息包括要在重建所述相邻段的所述步骤中使用的第一相邻段的指示。

第二段重建模式, 其中所述段重建信息包括运动系数信息。

50. 如权利要求 49 所述的方法, 其特征在于所述段重建模式组还
10 包括:

第三段重建模式, 其中所述段重建信息包括来自所述参考帧的象素值的指示;

第四段重建模式, 其中所述段重建信息包括来自所述当前帧的象素值的指示。

51. 如权利要求 49 所述的方法, 其特征在于所述第一相邻段的指示
15 包括有关所述第一相邻段相对于所述第二相邻段的位置的信息。

52. 如权利要求 51 所述的方法, 其特征在于所述第一相邻段的指示还包括有关所述第一相邻段中子段的信息。

53. 如权利要求 49 所述的方法, 其特征在于所述运动系数信息包
20 括具有至少一个非零运动系数值的指示。

54. 如权利要求 53 所述的方法, 其特征在于所述具有至少一个非零运动系数值的指示包括非零系数模式指示和至少一个非零系数值。

55. 如权利要求 49 所述的方法, 其特征在于所述第一段重建模式
25 包括利用从表示所述第一相邻段的第一运动场模型导出的预测运动场模型。

56. 如权利要求 55 所述的方法, 其特征在于通过将所述第一相邻段的所述第一运动场模型投影到所述第二相邻段来建立所述预测运动场模型。

00814037.5

权 利 要 求 书 第10/14页

57. 如权利要求 49 所述的方法, 其特征在于所述第二段重建模式包括利用求精运动场模型。

58. 如权利要求 57 所述的方法, 其特征在于所述求精运动场模型是通过至少一个运动系数值的指示来表示的。

5 59. 如权利要求 57 所述的方法, 其特征在于所述求精运动场模型表示第二运动场模型和所述预测运动场模型之间的差异, 其中所述第二运动场模型是从所述参考帧导出的所述第二段的表示。

60. 如权利要求 57 所述的方法, 其特征在于所述求精运动场模型是从所述参考帧导出的所述第二段的表示。

10 61. 在一种对视频序列进行编码的方法中, 所述视频序列包括至少一个当前视频帧和参考视频帧, 所述当前视频帧包括至少一个第一相邻段和第二相邻段, 一种用于当前视频帧的运动补偿预测的方法的改进包括:

为所述第二相邻段定义一组编码模式;

15 计算一组成本函数, 每一个所述成本函数与所述编码模式组中的一种模式相关联;

在所述成本函数组中选择具有最小绝对值的一个成本函数;

定义所述编码模式组中与所述最小绝对值相关联的一种编码模式作为用于所述第二相邻段的被选编码模式;

20 按照所述被选编码模式对有关所述第二相邻段的信息进行编码。

62. 如权利要求 61 所述的方法, 其特征在于还包括:

向解码器发送所述编码信息, 用于解码。

63. 如权利要求 61 所述的方法, 其特征在于还包括:

将所述编码信息储存在存储单元中。

25 64. 如权利要求 61 所述的方法, 其特征在于所述编码模式组包括:

第一编码模式, 其中来自所述第一相邻段的运动场模型被投影到所述第二相邻段以形成预测运动场模型, 并且所述第二相邻段由所述

00814037.5

权 利 要 求 书 第11/14页

预测运动场模型来表示;

第二编码模式, 其中所述第二相邻段由从所述参考帧导出的运动场模型来表示;

第三编码模式, 其中来自所述第一相邻段的运动场模型被投影到
5 所述第二相邻段以形成投影场模型, 并且所述第二相邻段由所述预测运动场模型和求精运动场模型来表示;

65. 如权利要求 64 所述的方法, 其特征在于所述编码模式组还包括:

第四编码模式, 其中利用来自所述参考帧的像素值对所述第二相
10 邻段进行编码;

第五编码模式, 其中利用来自所述当前帧的像素值对所述第二相邻段进行编码。

66. 如权利要求 64 所述的方法, 其特征在于所述求精运动场模型表示从所述参考帧导出的所述运动场模型和所述预测运动场模型之间
15 的差异。

67. 如权利要求 64 所述的方法, 其特征在于所述预测运动场模型、所述求精运动场模型和从所述参考帧导出的所述运动场模型包括一组基本函数, 每一个所述基本函数都与某个运动系数值相乘。

68. 如权利要求 67 所述的方法, 其特征在于所述基本函数是正交
20 函数。

69. 如权利要求 68 所述的方法, 其特征在于所述预测运动场模型、所述求精运动场模型和从所述参考帧导出的所述运动场模型是仿射运动场模型。

70. 如权利要求 61 所述的方法, 其特征在于所述成本函数组中的
25 每一个成本函数都包括遭受的图像失真的量度和在使用给定的一个所述编码模式时所需数据量的量度。

71. 如权利要求 70 所述的方法, 其特征在于利用拉格朗日准则来计算所述成本函数组中的每一个成本函数。

00814037.5

权 利 要 求 书 第12/14页

72. 如权利要求 71 所述的方法, 其特征在于所述拉格朗日准则具有 $L=D+\lambda \cdot B$ 的形式, 其中 D 是在对给定运动系数组进行编码时遭受的失真的量度, B 是表示所述运动系数所需的比特数, 而 λ 是拉格朗日参数。

5 73. 如权利要求 67 所述的方法, 其特征在于利用共同的基本函数组来表示所述预测运动场和所述求精运动场。

74. 如权利要求 67 所述的方法, 其特征在于通过去除运动系数来近似所述求精运动场模型。

75. 如权利要求 64 所述的方法, 其特征在于所述当前帧包括多个
10 第一相邻段, 所述方法还包括以下步骤:

形成多个预测运动场模型, 一个预测运动场模型对应于所述多个第一相邻段中各自的第一相邻段;

形成多个求精运动场模型, 各个求精运动场模型对应于所述多个预测运动场模型中给定的一个预测运动场模型。

15 76. 如权利要求 75 所述的方法, 其特征在于所述预测运动场模型是以一个以上第一相邻段为基础形成的。

77. 如权利要求 76 所述的方法, 其特征在于所述预测运动场模型是通过对来自一个以上的第一相邻段的运动场模型的投影求平均而形成的。

20 78. 如权利要求 64 所述的方法, 其特征在于所述方法还包括把所述第一相邻段分为多个子段并利用至少一个所述子段的运动场模型来形成所述预测运动场模型。

79. 如权利要求 61 所述的方法, 其特征在于所述对信息进行编码以依赖于所述被选场模型的方式而发生。

25 80. 如权利要求 79 所述的方法, 其特征在于: 如果所述被选编码模式是所述第二编码模式, 则所述方法还包括将所述求精运动场模型的所有运动系数设置为等于从所述参考帧导出的所述运动场模型的所述运动系数。

00814037.5

权 利 要 求 书 第13/14页

81. 如权利要求 80 所述的方法, 其特征在于所述对信息进行编码包括对所述求精运动场模型进行编码的步骤。

82. 如权利要求 79 所述的方法, 其特征在于: 如果所述被选编码模型是所述第一编码模式, 则所述对信息进行编码包括对所述预测运动场模型进行编码的步骤。

83. 如权利要求 79 所述的方法, 其特征在于: 如果所述被选编码模式是所述第三编码模式, 则所述对信息进行编码包括以下步骤:

对所述预测运动场模型进行编码;

对所述求精运动场模型进行编码。

84. 如权利要求 81 所述的方法, 其特征在于所述对所述求精运动场模型进行编码包括以下步骤:

通过对第一和第二值中供选择的一个设置运动系数指示符来指示所述编码信息包括所述求精场模型的所述运动系数;

通过设置运动系数指示符来指示所述求精场模型的所述运动系数中哪些具有非零值;

对所述非零值进行编码。

85. 如权利要求 83 所述的方法, 其特征在于所述对所述求精运动场模型进行编码包括以下步骤:

通过对第一和第二值中供选择的一个设置运动系数指示符来指示所述编码信息包括所述求精场模型的所述运动系数;

通过设置运动系数模式指示符来指示所述求精场模型的所述运动系数中的哪些具有非零值;

对所述非零值进行编码。

86. 如权利要求 84 所述的方法, 其特征在于各个所述非零系数值是通过指示幅度和符号来进行编码的。

87. 如权利要求 82 所述的方法, 其特征在于对所述预测运动场模型进行编码包括通过对第一和第二值中供选择的一个设置运动系数指示符来指示所述编码信息不包括运动系数值。

00814037.5

权 利 要 求 书 第14/14页

88. 如权利要求 87 所述的方法, 其特征在于对所述预测运动场模型进行编码还包括指示某个方向, 所述方向用于识别形成所述预测运动场模型的所述第一相邻段相对于所述第二相邻段的相对位置。

89. 如权利要求 88 所述的方法, 其特征在于对所述预测运动场模型进行编码还包括通过设置子段鉴别指示符来指示形成所述预测运动场模型的所述第一相邻段的子段。

90. 如权利要求 83 所述的方法, 其特征在于对所述预测运动场模型进行编码包括通过对第一和第二值中供选择的一个设置运动系数指示符来指示所述编码信息不包括运动系数值。

91. 如权利要求 90 所述的方法, 其特征在于对所述预测运动场模型进行编码还包括指示某个方向, 所述方向用于识别形成所述预测运动场模型的所述第一相邻段相对于所述第二相邻段的相对位置。

92. 如权利要求 91 所述的方法, 其特征在于对所述预测运动场模型进行编码还包括通过设置子段鉴别指示符来指示形成所述预测运动场模型的所述第一相邻段的子段。

00814037.5

说明书

第1/30页

用于压缩运动矢量场的装置和方法

5 本发明一般涉及对视频序列进行编码中利用运动补偿的方法。
更具体地说, 本发明涉及利用运动补偿预测对视频序列进行编码和
解码的装置以及相关方法。从视频帧的相邻段并通过利用正交仿射
运动矢量场 (orthogonal affine motion vector field) 模型来预测段的运
动场。通过本发明的一个实施例的操作, 运动矢量场可由数量减少
10 了的比特来构成, 同时仍保持低预测误差。

发明背景

数字通信技术中的进步已经允许发展新型和改进型通信类。更
进一步的进步将允许不断改进利用这种进步的通信和通信系统。

15 例如, 已经建议了用于能够形成视频帧的数字视频数据通信的
通信系统。电视会议期间利用的视频图像就是有利地利用了数字视
频序列的应用示例。

但是, 视频帧一般由大量的象素组成, 各个象素可用一组数字
比特表示。并且, 表示任何视频序列都需要大量的视频帧。由于形
20 成一个典型的视频序列所需的每帧中的大量象素以及大量的帧, 因
此表示视频序列所需的数据总量很快就变大。例如, 一示例性视频
帧包括 640×480 象素的阵列, 各象素具有 RGB (红、绿、蓝) 颜色,
其中各颜色分量用 8 比特表示, 每帧总共 7372800 比特。

视频序列, 就象记录在胶片上的普通运动图像一样, 包括一系
25 列静止图像, 形成运动的错觉是通过相对快速地显示连续图像, 比
如以每秒 15 至 30 帧的速率进行显示。由于相对快的帧速率, 连
续帧中的图像趋向于很相似。一般的景物包括一些例如背景景物的
静止组成部分和一些采用各种不同形式的运动组成部分, 所述运动

00814037.5

说明书 第2/30页

部分比如为新闻播音员的脸、移动的车辆等。另一方面，记录景物的摄像机本身是运动的，这时图像的所有组成部分具有相同类型的运动。在许多情况下，它表明一个视频帧和下一视频帧之间的总变化相当小。当然，这取决于运动的特性：运动越快，从一帧到下一帧的变化越大。

问题出现发送视频序列中，主要涉及必须从发送设备向接收机发送的信息量。序列的各帧包括矩形矩阵形式的象素阵列。为了得到清晰图像，需要高分辨率，也就是说，帧需要包括大量的象素。目前已有大量标准化图像格式，其中包括 352×288 象素的 CIF（公用中分辨率图像格式）和 176×144 象素的 QCIF（1/4 公用中分辨率图像格式）。QCIF 格式一般被用于第一代移动视频电话设备，并在可能用于这种设备的小型（3 至 4cm^2 ）LCD 显示器上提供可以接受的清晰图像。当然，较大的显示设备一般需要更高空间分辨率的图像，以便显示时这些图像能以足够的空间清晰度出现。

对图像的每一象素，必须提供颜色信息。如在上面提到的，颜色信息一般是按照基色分量红、绿和蓝（RGB）或利用相关亮度/色度模型进行编码，所述模型称为 YUV 模型，如下面将要描述的，它提供一些编码好处。尽管有几种可以提供颜色信息的方式，但所有颜色表示法都具有相同的问题，即正确表示出现在自然景色中的颜色范围所需的信息量的问题。为了建立人的视觉系统可接受质量的彩色图像，各个颜色分量一般必须用 8 比特分辨率表示。因此，图像的各个象素需要 24 个信息比特，这样 QCIF 分辨率的彩色图像就需要 $176 \times 144 \times (3 \times 8) = 608256$ 比特。此外，如果该 QCIF 图像构成每秒 15 帧的帧速率的视频序列的一部分，要编码该序列就需要总共 9123840 比特/秒。

照此，数据量有时必须通过相对低比特率的通信信道发送，诸如通过以低于 64kbps 工作的无线通信信道发送。

视频编码方案被用于减少表示这样的视频序列所需的数据量。

00814037.5

说明书 第3/30页

许多视频编码方案的核心是提供运动补偿预测的方法。运动补偿预测一般提供通过消除帧之间的时间冗余 (temporal redundancies) 来改进帧补偿的方法。对操作进行预测是基于下述事实：在短系列的相同一般图像 (the same general image) 中，大多数对象保持在相同位置，而其它对象仅移动短距离。这种运动作为二维运动矢量来描述。

利用 YUV 颜色模型可以获得一些编码优势。这利用了人的视觉系统的一种特性，即人的视觉系统对光强 (亮度) 变化较之颜色的变化更加敏感。因此，如果图像用一种亮度分量和两种色度分量 (如 YUV 模型中) 表示，有可能在空间上对色度分量进行二次抽样 (降低其分辨率)。这样作可以减少对图像中彩色信息进行编码所需的信息总量，具有可接受的图像质量降低。可有几种方式来执行空间二次抽样，但一般用表示亮度信息的一个 16×16 象素块和表示两种色度分量的一个 8×8 象素块来对图像中每一个 16×16 象素块进行编码。换句话说，在 x 和 y 方向按 2 的因数对色度分量进行二次抽样。所得到的一个 16×16 亮度块和两个 8×8 色度块的组合通常被称为宏块。利用这种编码方案，对 QCIF 图像进行编码所需的信息总量计算如下：QCIF 分辨率是 176×144 。所以图像中包含 11×9 个 16×16 象素亮度块。每个亮度块具有两个与之相关联的 8×8 象素的二次抽样色度块，也就是图像中有 11×9 个宏块。如果亮度和色度分量用 8 比特分辨率编码，则每个宏块所需的总比特数是 $1 \times (16 \times 16 \times 8) + 2 \times (8 \times 8 \times 8) = 3072$ 比特。对整个 QCIF 图像进行编码所需的比特数现在是 $99 \times 3072 = 304128$ 比特，即是没有执行色度子抽样所需的数量的一半 (见上述内容)。但是，这仍是非常大的信息量，如果以这种方式编码的 QCIF 图像是每秒 15 帧视频序列的一部分，则仍需要总共 4561920 比特/秒。

视频编码需要处理大量的信息。这当然意味着对视频图像进行编码需要强大的信号处理设备，如果这些图像以其原形式发送，则

00814037.5

说明书 第4/30页

需要宽带宽的通信信道。但是，在许多场合不可能提供高容量的传输信道。特别是在电视电话应用中，视频信号必须通过现有的固定线路通信信道（即通过传统的公共电话网）发送或利用诸如移动电话网提供的无线电通信链路发送，情况更是如此。现在已有许多为
5 这些系统中的视频编码制定准则的国际电信标准。国际电信联盟（ITU）标准的 H.261 和 H.263 就是示例。标准 H.261 给出了以 64kbps 的倍数工作的传输系统（这些一般是固定线路电话网）中视频编码的建议，而 H.263 对可用带宽小于 64kbps 的系统提供了类似的建议。这两个标准实际上关系非常密切，并且都利用能够减少必须传送的信息量的运动预测编码技术。
10

在移动电视电话中，目的是通过可用带宽约 20kbps 的传输信道发送视频序列。典型的帧速率应该足够提供良好的运动错觉，因此应该在每秒 10 和 15 帧之间。因此应该意识到，为了将需要约 4.5Mbps 的视频序列与只能传送 20kbps 的信道相匹配，需要非常大的压缩率
15 （约 225:1）。这就是运动预测编码以及其它技术开始起作用的情况。

运动预测编码的基本思路是考虑到视频序列中存在的大量时间冗余。如上面所解释的，在以比较快的帧速率（即大于每秒 10 帧）记录的典型视频序列中，从一帧到下一帧只有很小的变化。通常背景是静止的，只有图像的某些部分经历某种形式的运动。另一方面，
20 如果照相机本身在动，则所有组成部分都经历某种一致的运动。

因此在尝试减少发送视频序列时的信息量的时候，可利用连续帧之间的高度相关性。换句话说，一帧可以从之前的所谓参考参考帧来预测通常是直接在当前被编码的帧之前的帧，但也不一定。在这样一种编码方案中，一般只有当前帧和参考帧之间的不同之处被
25 编码并被发送到接收机。一般来说，这种编码被称为 INTER 编码。这种编码方案必须要求发射机和接收机都保留参考帧（例如之前编码的帧）的记录。在发射机中，视频编码器对当前帧和参考帧进行比较，识别两帧之间的差异，将之编码，并向接收机传送有关变化

00814037.5

说明书 第5/30页

的信息。在接收机中，通过将所述两帧之间的差异信息加到参考（如之前）帧中而在视频解码器中重建当前帧。储存在编码器和解码器中的帧接着被更新，这样当前帧就成为新的参考，并继续以相同的方式从一帧处理到下一帧。

5 当然有一些场合不能使用这种预测。显然，视频序列的第一帧总是必须被编码并被发送到接收机中的解码器。显然也没有之前的帧可以用作预测编码的参考。相似的情况发生在景物剪辑的情形。这时当前帧可能和之前的帧非常不同，以致没有可能预测，因此，新的帧必须被编码和发送。这种编码一般被称为 INTRA 编码。许多
10 编码方案也使用周期性 INTRA 帧编码。例如，每 10 或 20 帧发送一个 INTRA 帧。这样做可以中和编码错误的影响，编码错误可能逐渐累积，最终导致重建图像中不可接受的失真。

运动预测编码可以被看作是上面介绍的 INTER 编码技术的扩展。上述内容描述了差异信息如何被发送到接收机，从而能够参考
15 某个之前的帧来对当前视频帧进行解码。提供差异信息的最简单又最明显的方式将是发送当前图像中与参考图像中相应的象素不同的各个象素的象素值（YUV 数据）。但是，实际上这种解决方案没有提供通过非常低比特率的信道进行视频传输所必要的数据率的减少。运动预测编码采用不同的途径。如上所述，编码器和解码器都
20 保留参考帧的记录，并参考所储存的帧对当前帧进行编码。在解码器中，参考所储存的之前的帧和编码器发送的差异信息来重建当前图像。

在编码器中，为了确定当前帧和参考帧之间的对应关系，一段一段地对当前帧进行检查。可以采用各种分段方案。常常是，简单
25 地将当前图像分割为规则的象素块。例如比较可以从一个宏块到一个宏块地进行。要不可按其它方式分割帧；或许试图更好地识别帧中包含的差异象素并能够更准确地确定帧中的运动。

利用预定的分段方案，在当前帧的各段和参考帧之间进行比较，

00814037.5

说明 15 第6/30页

以便确定该段中的象素和参考帧中的某组象素之间的“最佳匹配”。注意到对参考帧没有固定的分段；最佳对应于当前帧的给定段的象素在某些限制条件下可在参考帧中的任何位置，所述某些限制在下面会被解释。这样，运动预测编码可被视为识别当前图像中一组象素的原点的尝试，即试图通过回顾参考帧来确定象素值是如何从一帧到下一帧传播。

一旦找到当前帧中给定段的最佳匹配，就利用“运动矢量”对段和参考帧之间的对应关系进行编码。运动矢量可以被看作是 x 和 y （水平和垂直）分量的位移矢量，它实际上从当前帧的段指回参考帧中的象素位置。因此运动矢量实际上通过与参考帧进行比较识别当前帧中象素的原点。编码一直进行到当前帧中各段的原点都被识别为止。所得的表示可以被认为是描述两个帧之间整体对应关系的“运动矢量场”。

利用运动矢量一段一段地编码完整的视频帧产生对当前帧非常有效的表示，因为有关各段运动矢量的编码信息需要相对非常少的比特。但是，编码过程并不完美，存在错误和信息丢失。通常，错误的发生是由于不能正确识别参考帧中的对应象素值。例如，可能存在从一帧到下一帧的图像内容的某种变化，因此在当前帧中出现参考帧中没有对应物的新象素。此外，许多预测运动编码器限制了帧之间允许的运动类型。这种限制发生如下：为了进一步减少表示运动矢量场所需的信息量，运动预测编码器一般使用“运动模型”描述象素值从一帧到下一帧传播的方式。利用运动模型，依据一组“基本函数”来描述运动矢量场。依据这些数学基本函数来表示从一帧到下一帧的象素值传播是。通常，运动被表示为涉及这些基本函数与某些系数值相乘的和，按为提供运动矢量场的最佳近似的这样一种方式来确定这些系数。运动矢量场的这种重新表达式（re-expression）必然引入某种附加误差，因为运动模型不能准确描述运动矢量场。但是，这种方法有一个显著的优点，因为现在只有运动

00814037.5

说明书 第7/30页

模型系数必须被发送到解码器。这个优点的出现，因为按照所需精确度的实现和程度，预先选择运动场基本函数，并且使得编码器和解码器都知道这些函数。许多目前建议的视频编码方案利用运动预测编码，具体地说，H.263 标准，这些视频编码方案是基于平移运动场模型，即该模型的基本函数只能够表示 x 和 y（水平和垂直）方向中的直线运动。因此无法表示连续帧之间可能发生的像素旋转和歪斜，这不可避免地在预测运动中引入误差。

5 最后，为了补偿运动矢量场编码处理引入的误差，典型的运动预测编码器包括误差估计功能。有关预测误差的信息和运动场模型系数一起被发送到解码器。为了估计运动场编码处理引入的误差，运动预测编码器一般也包括解码部分，该解码部分与接收机中的一样。一旦利用上述运动预测方法编码了当前帧，编码器的解码部分就重建当前帧，并将其与当前帧的原始版本相比较。于是就有可能建立“预测误差帧”，其中包含编码当前帧和原始当前帧之间的不同。这个信息连同运动场模型系数、或许还有有关当前帧分段的信息一起被发送到解码器。

即使利用了这样的示例，表示视频序列仍然需要大量数据。

利用减少了的比特量或降低了的比特率对视频序列进行编码而同时保持低预测误差的改进方法将因此是有利的。

20 按照与视频数据相关的背景信息，发展了本发明的重大改进。

发明概述

因此，本发明有利地提供了用于利用运动补偿预测来操作视频序列的装置和相关方法。

25 提供一种方法，通过将视频帧分成段并从某个段的相邻段来预测该段的运动场、以及通过利用正交仿射运动矢量场模型，来表示运动矢量场。本发明的一个实施例的操作提供了一种快速简洁对运动矢量场进行编码而同时仍保持低预测误差的方式。由此提供质量

00814037.5

说明书 第8/30页

改进的共同构成视频序列的视频帧的通信。

通过本发明的实施例的操作，提供一种减少表示运动矢量场所需信息量而同时保持低预测误差量的方式。

提供一种形成运动矢量场的用于编码器的运动场编码器。应用
5 由仿射运动矢量场模型构成。相反，例如，对于纯粹的平移运动模型，通过利用仿射模型可以获得运动场的更灵活的表达式。诸如变焦、旋转、转向（sheer）、或平移的典型自然运动不能由仿射运动矢量场模型来表示。仅仅利用平移模型的传统系统不能表示其它形式的运动。

10 视频帧的相邻段的仿射运动矢量场的相似性通过利用仿射预测运动矢量场而被利用。如果，例如，两个相邻段具有相似的运动矢量场，则其中一个运动矢量场能通过另一个运动矢量场仅外加一个小的、或甚至忽略不计的、也就是零的求精（refinement）场来算得。对于视频帧的各个段，选择一种仿射运动模型，它利用尽可能少的
15 非零系数达到令人满意的低预测误差。此外，利用正交基本函数。正交基本函数对相应运动系数的量化敏感度低，这样系数就能用少量比特来表示。那就是说，运动系数的有效传输需要要被量化到低精确程度的系数。但是，在用少量比特进行表示时，传统使用的基本函数的类型导致不可接受的预测误差增加。由于对应于正交基本
20 函数的系数对量化更加健壮，因此在本发明实施例的操作中有利地利用了正交基本函数。

在本发明的一个方面，对视频编码器提供运动场编码器。运动场编码器可用于形成压缩运动矢量场，所述压缩运动矢量场由一组
25 当前帧的所有象素的运动矢量构成。运动矢量场由预测运动矢量场和求精运动矢量场构成。

在本发明的另一个方面，对视频编码器提供了运动补偿预测装置，运动补偿预测装置接收由运动场编码器形成的压缩运动矢量场的指示。运动补偿预测装置建立预测帧。预测装置能够通过计算帧

00814037.5

说明书 第9/30页

的各个段的运动矢量场来重建帧的象素。根据预测运动矢量场和求精运动矢量场来计算运动矢量场。

在本发明的又一个方面，对视频解码器提供运动补偿预测装置。运动补偿预测装置接收预测运动矢量场和求精运动矢量场系数的指示。

因此，在这些和其它方面，提供用于操作视频序列的视频设备的装置。视频序列至少包括当前视频帧，所述当前视频帧至少具有第一相邻段和第二相邻段。所述装置形成第二相邻段的运动矢量场的近似。所述装置包括运动矢量场建立器，它被连接用来接收表示第一仿射运动模型的指示，形成表示第一相邻段的第一运动矢量场的近似。运动矢量场建立器对表示第一仿射运动模型的指示进行响应，形成第二仿射运动模型。第二仿射运动模型形成第二相邻段的运动矢量场的近似。

通过下面简要描述的附图、本发明优选实施例的以下详细描述及后附权利要求书，可以更完整了解本发明及其范围。

附图说明

图 1 示出本发明一个实施例可工作的视频通信系统的编码器和解码器。

图 2 示出运动场编码器的功能方框图，它构成图 1 所示通信系统的一部分。

图 3 示出运动补偿预测装置的功能方框图，它构成图 1 所示通信系统的编码器还有解码器的一部分。

图 4 示出在本发明一个实施例的操作期间把视频帧分为段的方式。

图 5 中的表表示在本发明的实施例的操作期间利用的选择比特的示例值和意义。

00814037.5

说明书 第10/30页

详细描述

本发明报告的一个实施例的运动预测视频编码的新方法进一步减少了在低比特率视频编码系统中从编码器向解码器传送的数据量，同时保持良好的图像质量。该方法包括新方法，它利用相同帧中已编码的段进一步预测当前帧中段的象素值。

在一个示例性实施例中，当新的视频序列要被编码和发送时，序列中的第一帧以 INTRA 格式被发送，象先有技术和上面描述的那样。所述帧接着被储存在编码器和解码器中，并构成序列中下一（即第二）帧的参考帧。

当编码器开始对第二帧进行编码时，它通过检查帧的第一段来开始编码处理。在优选实施例中，当前帧被分成一组 16×16 象素段，但这不是所述方法的根本，可以设想其它分段方案。编码从上面最左段开始，并从左到右和上到下进行到整个帧（即编码处理按行从上到下进行）。

确定描述参考帧和当前帧的第一段之间的象素值的映射的运动矢量场，接着使用所谓的“仿射”运动模型来近似运动矢量并生成一组运动系数。仿射运动模型是一种特殊类别的运动模型，其数学形式诸如允许帧之间的平移、旋转和歪斜运动。它包括 6 个基本函数。因此运动矢量基本上由涉及六个基本函数与适当选择的“运动系数”相乘的和来代替。于是仅仅向解码器发送运动系数（或它的子集）就足够了，因为编码器和解码器两边都知道基本函数本身。对当前帧的任何给定段产生最佳仿射运动模型的参考帧中的象素组可以驻留在参考帧的任何区域中，至少理论上是这样。这里应该强调的是，此方法的目的在于不仅仅是使预测误差最小，也是为了找到产生“率失真（rate-distortion）”意义上的段的最佳匹配的仿射运动场模型。这意味着最佳匹配是通过考虑图像失真量度和达到所述失真度（level of distortion）所需数据量的量度两者来确定的。

由于帧的第一（上面最左边）段没有之前编码的相邻段，不能

00814037.5

说明书 第11/30页

采取进一步的动作，编码器进行到当前帧的第二段。接着确定在参考帧和当前帧的第二段之间提供最佳映射的仿射运动场模型，其中利用与之前描述的相同种类的率失真、最佳匹配估计。如前所述，象素值的相应区可以驻留在参考帧的任何地方，并且实际上可以与之前被确定为当前帧第一段的最佳匹配的部分相重叠。

第二段具有一个之前编码的相邻段（即第一段）。编码器现在考虑根据之前为第一段确定的仿射运动模型来模拟第二段是否比根据为第二段本身新近确定的仿射运动系数“更有效”。基本原理如下：由于用于第一段的运动系数已被确定并发送到解码器，有可能在编码第二段的时候减少必须发送到解码器的信息量。因此提高编码效率。

但是，用于第一段的运动系数不太可能正好与那些最准确模拟第二段的运动矢量场的运动系数一样。所以，不是简单地那样利用为第一段计算的运动系数，而是执行一种投影，以便将第一段的运动场映射到第二段。即使执行了这种投影，有关第一和第二段运动场之间差异的一些信息仍有可能必须被发送到解码器，以避免在重建图像中的不可接受的失真。因此，编码器比较 a) 发送特别为第二段确定的运动系数数据所需的数据量和 b) 在第二段的运动矢量场是从第一段的运动模型的投影加上一些“求精”信息来确定时所需的数据量。在选择发送什么信息时，编码器也必须考虑由预测处理引入到图像的失真。这种选项之间的比较可以被认为是确定选择特定选项的“成本”，即要发送的信息量和允许的失真量之间的一种折衷。

这种方法对运动预测编码的好处可能不会马上表现出来。但是，在许多情况中发现，在从相邻段投影运动场模型之后，需要非常少的、或者甚至为 0 的求精信息。这一点能够导致必须从编码器发送到解码器的数据量的显著减少。在需要 0 求精信息的情况下，第二段的运动矢量场可以仅仅根据已经储存在解码器中的运动系数来预

00814037.5

说明书 第12/30页

测。

此范例到目前为止，仅仅考虑了帧的第一和第二段。如上所述，按照本发明优选实施例中使用分段方案，第二段只有一个能被用于预测其运动系数的相邻段。这一点对帧的第一行上其它所有段都一样。所有这样的段只能具有紧靠其左边的之前编码的相邻段。但是，在图像的第二及后续行，在各行上面的之前编码的段也可以用。因此，后续行的段具有左边和上面的相邻段。这一点对除了各行中的第一段的所有其它段都适用，各行中的第一段仅仅只有直接在其上的之前编码的相邻段。所以，当考虑帧中要编码的一般段时，对运动系数的预测存在几种可能性。一般情况下，编码器可以试着采用给定段的上面段或左边段的运动场模型来预测该给定段的运动系数。另一方面，可利用两个相邻段的运动场模型来形成某种平均。在各种情况下，从相邻段预测的运动场模型被称为“预测场”，而预测场和特别对给定段本身确定的运动场模型之间的差异被称为“求精场”。在优选实施例中，预测场和求精场都是仿射运动场模型。预测场和求精场之和应等于通过将仿射运动模型应用到给定段本身所确定的运动场模型。在不可能从给定段的相邻段预测该给定段的运动场模型的情况下，预测场被设为零，而求精场等于特别对给定段本身确定的运动场模型。

从上面的描述将认识到，给定段能以几种不同的方式被编码。在编码器中利用那个选项的选择是基于之前描述的“率失真”的考虑来作出的。因此，根据选择的编码选项，有几种不同类型的数据必须被发送到解码器，所述信息必须以清楚的方式被发送到解码器，从而使段能被正确重建和显示。各种编码选项如下：1.) 给定段可以被表示为预测场和求精场之和。2.) 给定段可以仅被表示为预测场。这种情况在给定段能适当根据一个或多个之前被编码的相邻段的运动场来表示，而不需要求精信息的时候，或者编码器已经发现将求精场减少到零有效的情况下发生。3.) 有问题的段可以利用通过参考

00814037.5

说明书 第13/30页

帧特别对给定段确定的运动模型来进行编码。在这种情况下，如上所述，预测场被设为零，而求精场被设为等于从参考帧确定的运动场模型。

基本上，为了能够正确重建给定段，有两类必须发送到解码器的信息。它们是：1) 选择信息，它使解码器能够选择正确的相邻段在预测中使用；2) 运动系数信息。每当利用预测场对段进行编码时，不论是否存在相关的求精场，都有必要提供有关用于预测的相邻段的信息。不必发送任何运动系数数据，因为解码器已经知道（即储存在其中）之前编码的相邻段的运动场模型。如果例如预测是基于一个以上相邻段、或相邻段已被分为子段而一个或多个子段的运动场模型被用于形成预测场，则也可能需要附加信息。在使用求精场时，必须提供运动系数值。在这种情况下，应该记住只需要发送运动系数数据，因为解码器和编码器都知道运动模型基本函数。

从编码器发送到解码器的数据流因此可能包含运动系数数据和各种指示解码器执行不同操作的选择数据（即，非运动系数数据）。例如，如果解码器接收到非运动系数数据，则它应该利用选择数据指示的相邻段和子段来建立预测运动场模型。如果它接收到运动系数数据，则解码器必须利用发送的运动系数值和储存的运动模型基本函数来建立求精运动场模型。在本发明优选实施例中由编码器提供的数据流的格式将在下文中进行详细描述。

其它一些求精方法都是可能的。在本发明的优选实施例中，相邻段可以被分为更小的子段。具体地说，各个 16×16 像素段可以被分为四个 8×8 像素块，这些子段的运动场模型可被用于得到预测场。在这种情况下，一个一般的 16×16 像素段具有四个紧邻的 8×8 像素子段，它们可被认为是两个正好在上面和两个紧靠左边。在这种情况下，决定处理更复杂了些，但与前面段落中描述的操作方式仍基本相同。子段大小的选择不限于刚才介绍的例子，可以考虑各种其它子段的大小。例如， 4×8 或 8×4 像素块都可以被用作子段。

00814037.5

说明书 第14/30页

如上所述，在按照本发明的方法应用于实际中时，常常发现需要几乎很少的求精信息，一般段的运动模型能以非常高的精度从其相邻段的运动场模型来预测。本发明还包括一个特性，求精场或整个求精场的各个系数可被设为零，如果这在“率失真意义”上有效的话。换句话说，在考虑到减少要发送的数据量时，求精场可被设为 0，如果这样做引入的图像失真可以接受。这个附加特性进一步减少了必须从编码器发送到解码器的数据量。

首先参考图 1，一般用 10 表示的一种通信系统，可用于视频序列发生器和视频序列接收机之间的视频序列通信。图中示出视频序列发生器的编码器 12，并示出构成视频序列接收机的一部分的解码器 14。为了简单起见，没有示出视频序列发生器和接收机上各自的其它部件。通信路径 16 被示为对通信系统各部分的互连。通信路径可以采用任何形式，包括例如无线电链路。

这里示出编码器 12 被连接用于接收线路 18 上的视频输入。视频输入被提供给运动估计器 22 和减法器 24 的输入端。运动估计器也被连接用于接收储存在帧存储器 26 中的参考帧的指示。运动估计器计算正被编码的帧和之前的帧之间的像素的运动矢量，即当前视频输入 $I_n(x, y)$ 和参考帧 $R_{ref}(x, y)$ 之间的像素的运动矢量。

一旦编码器对各个段进行了编码，重建所需的信息就可以被发送到解码器，并且解码器就可以开始重建所述段。因为各帧是按一段一段的原则编码的，并且预测处理中只使用之前编码的段，因此解码器中帧的重建可以立即开始，即不必等到整个帧都被编码。一能得到有关各段的信息，就将该信息发送到解码器，并在接收机中基本与编码处理同时地对帧进行解码。在电视电话应用中，使端到端的延迟保持最小是有利的。当然，该方法也可适用于视频存储和检索系统，其中即时传输不是必须要求的。在那种情况下，不要求即时发送数据并且也可能在当前帧中使用其它相邻段用于预测目的。

00814037.5

说明书 第15/30页

运动估计器 22 被连接到运动场编码器 28。运动场编码器 28 能用于形成一个运动矢量场，该运动矢量场是当前帧的所有象素的一组运动矢量。运动场编码器产生的场通过线路 32 提供给复用器 34，其后通过通信路径 16 被传送到视频序列接收机及其解码器 14。

5 所示编码器还包括运动补偿 (MC) 预测装置 36。预测装置 36 也被连接到帧存储器 26。预测装置 36 能用于产生预测帧，所述预测帧被提供给减法器 24 和加法器 38。

减法器 24 形成的差值被提供给预测误差编码器 42。该预测误差编码器确定当前输入视频帧和该帧的 MC 预测版本之间的象素值的差异，以便产生预测误差的指示。并且，预测误差编码器 42 又被连接到复用器 34 以及预测误差解码器 46。预测误差解码单元对预测误差进行解码，所述预测误差由加法器 38 加到 MC 预测当前帧，所得结果储存在帧存储器 26 中。

这里示出的解码器 14 包括分用器 52、预测误差解码器 54、运动补偿预测装置 36、加法器 56 以及帧存储器 26。编码器和解码器的预测装置 36 以及各个设备中的帧存储器 26 用相同的数字表示。

运动估计器 22 计算正被编码的称为当前帧 $I_n(x, y)$ 的帧和参考帧 $R_{ref}(x, y)$ 之间的象素的运动矢量 $(\Delta x(x, y), \Delta y(x, y))$ 。参考帧是之前被编码和发送的帧之一，在给定瞬间在编码器和解码器的帧存储器 26 中可以得到。

$\Delta x(x, y)$ 和 $\Delta y(x, y)$ 分别是水平和垂直位移值。当前帧中被称为运动矢量场的所有象素的运动矢量组通过运动场编码器 28 被压缩，其后，如上所述，被发送到解码器。

25 为了指示运动矢量场的压缩一般是有损耗的，压缩的运动矢量被表示为 $(\Sigma x(x, y), \Sigma y(x, y))$ 。在运动补偿预测装置 36 中，压缩的运动矢量和参考帧被用于建立预测帧 $P_n(x, y)$ 。预测帧是利用由运动估计器 22 和运动场编码器 28 以及参考帧 $R_{ref}(x, y)$ 的象素值确定的运动矢量场来计算的当前帧 $I_n(x, y)$ 的编码版本。

00814037.5

说明书 第16/30页

下列公式示出了计算预测帧的方法：

公式 1

$$P_n(x, y) = R_{ref}(x + \bar{\Delta}x(x, y), y + \bar{\Delta}y(x, y))$$

5 预测误差，即当前帧和预测帧之间的差异，如下：

公式 2

$$E_n(x, y) = I_n(x, y) - P_n(x, y)$$

10 预测误差被压缩并发送到解码器 14。压缩的预测误差被表示为 $\tilde{E}_n(x, y)$ 。

在解码器 14 中，当前编码的帧 $I_n(x, y)$ 的象素通过利用接收的运动矢量在参考帧 $R_{ref}(x, y)$ 中找到预测象素并通过按如下公式加上接收的预测误差 $\tilde{E}_n(x, y)$ 而被重建：

公式 3

15

$$I_n(x, y) = R_{ref}(x + \bar{\Delta}x(x, y), y + \bar{\Delta}y(x, y)) + \tilde{E}_n(x, y)$$

编码帧和原始帧之间的差异用如下公式表示，并且被称为重建误差：

公式 4

20

$$D_n(x, y) = I_n(x, y) - \tilde{I}_n(x, y)$$

由 MC 预测装置 36 形成的运动补偿预测帧 $P_n(x, y)$ 以这样一种方式被建立：使重建误差量最小，同时使表示运动矢量场所需的信息量最小。

00814037.6

说明书 第17/30页

典型视频序列的帧包含多个具有不同运动的段。所以，运动补偿预测的执行是通过将帧 $I_n(x, y)$ 划分为若干段并估计这样的帧和参考帧之间这样的段的运动。分段信息是运动表达式的固有部分。除非使用缺省帧分段，并且为编码器和解码器所知，否则描述帧的最终划分的附加信息必须被发送给解码器。实际上，一个段一般包括至少几十个象素。为了简洁表示这样的象素的运动矢量，它们的值最好用具有几个参数的函数来描述。这样的函数被称为运动矢量场模型。为了下面的描述，图像段的运动矢量应利用下面一般的加法表达式近似：

10 公式 5

$$\Delta x(x, y) = \Delta x_{prd}(x, y) + \Delta x_{refine}(x, y) \quad \Delta y(x, y) = \Delta y_{prd}(x, y) + \Delta y_{refine}(x, y)$$

上述公式的第二项被称为求精运动矢量场，并表示为如下线性组合：

15 公式 6

$$\Delta x_{refine}(x, y) = \sum_{n=1}^N c_n f_n(x, y) \quad \Delta y_{refine}(x, y) = \sum_{n=N+1}^{N+M} c_n f_n(x, y)$$

20 参数 c_n 被称为求精运动系数。所述系数在编码器中被压缩、通过通信路径 16 被发送，然后在解码器中被恢复。

函数 f_n 被称为基本函数，并且为编码器 12 和解码器 14 所知。矢量组 $(\Delta x_{prd}(x, y), \Delta y_{prd}(x, y))$ 被称为预测运动矢量场，并为编码器 12 和解码器 14 所知。

25 运动补偿预测之后得到的公式 2 中的预测误差帧 $E_n(x, y)$ 一般利用诸如离散余弦变换 (DCT) 的二维变换进行编码。这个处理被称为预测误差编码，其目的是减少预测误差。由于预测误差编码一般有损耗，这种导致重建误差。

编码器 12 的主要任务是找到适当的运动系数组，所述运动系数组要被编码并发送到解码器。通常，通过增加分配给编码系数的比

00814037.5

说明书 第18/30页

特数来减少作为结果的所遭受的失真。但是，失真的减少并非总是值得增加比特数。通常，处理这样的折衷的一种方式使下面的拉格朗日准则 (Lagrangian criterion) 最小：

公式 7

$$L = D + \lambda \cdot B$$

5

在这个公式中，项 D 表示在由给定的系数组进行编码时遭受的失真，即误差。发送系数的成本用比特数 B 表示。因子 λ 是称为拉格朗日参数的恒量。

在本发明一个实施例的操作中，视频帧的给定段的运动矢量场是两个仿射运动矢量场之和，即预测运动矢量场和求精运动矢量场之和，用如下公式表示：

公式 8

$$\Delta x(x, y) = \Delta x_{\text{pred}}(x, y) + \Delta x_{\text{refine}}(x, y) \quad \Delta y(x, y) = \Delta y_{\text{pred}}(x, y) + \Delta y_{\text{refine}}(x, y)$$

15 预测运动矢量场是以几种方式中的其中一种方式从一个或多个相邻段的运动矢量场得到的。例如，在一个实现中，预测运动矢量场是通过将例如邻近段的相邻段的仿射运动矢量场外插到当前段覆盖的区域得到的。由于当前段可以有几个相邻段，一般向解码器提供信令信息 (signaling information)，以便指定将使用那一段。在另一个实现中，预测运动矢量场是利用编码器和解码器已知的某种特定方法来组合几个相邻段的仿射运动矢量场而得到的。这样的方法是例如对水平和垂直运动矢量场分量求平均或确定中值。

求精运动矢量场具有按如下公式表示的仿射模型。

公式 9

25

$$\Delta x_{\text{refine}}(x, y) = \sum_{n=1}^3 c_n f_n(x, y) \quad \Delta y_{\text{refine}}(x, y) = \sum_{n=1}^3 c_{n+3} f_n(x, y)$$

00814037.5

说明书 第19/30页

其中基本函数 f_1, \dots, f_3 是仿射正交函数。基本函数关于确定给定段界限的矩形是正交的。同时，系数 c_1, \dots, c_6 是对应于正交基本函数组的求精运动矢量场系数。

- 5 由运动场编码器在编码器 12 进行编码期间对帧中的每一个段确定求精运动系数，具体地说是由运动场编码器 28 进行。

图 2 更详细地说明了运动场编码器 28。这里示出的编码器 28 包括预测运动场的选择器和建立器 62、运动分析器 64、运动系数取消装置 66 及量化器 68。

- 10 选择器和建立器 62 能用于给定段，以确定当前帧的之前编码段、或这样的段的组合，其运动矢量场最适合预测例如当前段的给定段的运动场。根据“胜出”候选的运动矢量场，如上所述地计算预测运动场。通常，信令信息被发送给解码器以指定几个候选段中最适合的那个。

- 15 运动分析器 64 能用于找到求精运动矢量场的新表达式。那就是说，得到一种数学上有效的表达式。新表达式随后被用于运动系数取消装置 66 中快速而灵活确定求精运动系数。

运动系数取消装置 66 能用于确定应设为 0 的求精系数，并计算剩下的非零系数的值，从而使拉格朗日准则最小，如下式：

- 20 公式 10

$$L(c) = D(c) + \lambda \cdot B(c)$$

- 其中 $D(c)$ 和 $B(c)$ 是与利用求精运动系数 c 对给定段进行编码相对应的预测误差和比特的量度，常数 λ 是拉格朗日参数。当设置某些求精运动矢量场系数为 0 时，预测误差增加。但是，当更多的系数被设为 0 时，编码器需要向解码器发送的比特数减少。所以，拉格朗日的值在某些求精运动系数被设为 0 时可以降低。

量化器 68 能用于量化剩余的非零求精运动矢量系数，以便使这样的系数适合于熵编码及从编码器向解码器的传输。

00814037.5

说明书 第20/30页

图 3 示出运动补偿 (MC) 预测装置 36 分别形成如图 1 所示的编码器和解码器 12 和 14 的一部分。编码器和解码器的 MC 预测装置 36 的功能单元相似, 并且 MC 预测装置能用在编码器和解码器中通过计算帧内各段的运动矢量场来重建当前帧的象素。根据预测运动矢量场 ($\Delta x_{prd}(x, y)$, $\Delta y_{prd}(x, y)$) 和求精运动矢量场系数来计算运动矢量场。在示例性实施例中, 求精运动矢量场由其反量化值表示。在解码器 14 中, 从一个或几个已解码的相邻段来得到预测运动矢量场。在解码和反量化器 76 执行反量化之后, 在解码器中可以得到求精运动矢量场系数。如图所示, MC 预测装置还包括运动矢量场建立器、段预测装置 80 和预测运动矢量场建立器 81。

如图 2 所示, 到运动场编码器 62 的运动分析器 64 的输入包括估计运动矢量场 ($\Delta x(x, y)$, $\Delta y(x, y)$)。运动矢量场由运动估计器 22 (图 1 所示) 提供。以传统的方式在运动估计器 22 中计算运动矢量场。预测运动矢量场也被提供给运动分析器。并且, 要编码的段 S 的几何形状及参考和当前帧 (分别为 $R_{ref}(x, y)$ 和 $I_n(x, y)$) 也都作为输入提供给运动分析器, 其中所述给定段 S 的几何形状也就是给定段 S 的大小和形状。

运动分析器能用于执行几种操作。首先, 运动分析器执行误差线性化。给定段 S_i 的预测误差 D_i 用公式 11 表示, D_i 包括 P 个象素坐标 (x_p, y_p), $p = 1, 2 \dots P$ 组成, 其预测运动场表示为 ($\Delta x_{prd}(x_p, y_p)$, $\Delta y_{prd}(x_p, y_p)$) 并且其求精运动矢量场由公式 9 给出的仿射运动模型近似:

公式 11

$$D_i = \sum_{p=1}^P (I_n(x_p, y_p) - R_{ref}(x_p + \Delta x_{prd}(x_p, y_p) + \Delta x_{refine}(x_p, y_p), y_p + \Delta y_{prd}(x_p, y_p) + \Delta y_{refine}(x_p, y_p)))$$

25

在线性化期间, 公式 11 的 $R_{ref}(x, y)$ 值利用某种已知近似方法来近似, 以便它变得线性依赖于 ($\Delta x_{refine}(x_p, y_p)$, $\Delta y_{refine}(x_p,$

00814037.5

说明书 第21/30页

$y_p)$)。于是,平方预测误差 D_i 可以近似如下:

公式 12

$$D_i = \sum_{p=1}^P (e_{p,1}c_1 + e_{p,2}c_2 + \dots + e_{p,6}c_6 - w_p)^2$$

5 e 和 w 的值取决于使用的近似方法的类型。

其后,利用运动分析器建立矩阵。由于在公式(12)中的平方下的元素是系数 c_n 的线性组合,因此公式的最小化正好等于下列矩阵表达式的最小化。

公式 13

10

$$(E_i c_i - w_i)^T (E_i c_i - w_i)$$

其中 E_i 、 w_i 和 c_i 下:

公式 14

$$E_i = \begin{bmatrix} e_{1,1} & e_{1,2} & \dots & e_{1,N+M} \\ e_{2,1} & e_{2,2} & \dots & e_{2,N+M} \\ \dots & \dots & \dots & \dots \\ e_{p,1} & e_{p,2} & \dots & e_{p,N+M} \end{bmatrix}, w_i = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_P \end{bmatrix}, c_i = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_{N+M} \end{bmatrix}$$

15

根据 E_i 和 w_i , 矩阵 A_i 和矢量 d_i 计算如下:

公式 15

$$A_i = E_i^T E_i.$$

20

公式 16

$$d_i = E_i^T w_i$$

00814037.5

说明书 第22/30页

运动分析器产生一个包括 $(N + M) \times (N + M)$ 的上三角形矩阵 R_i 的输出。所述 R_i 具有如下形式：

$$R_i = \begin{bmatrix} x & x & x & \cdots & x \\ 0 & x & x & \cdots & x \\ 0 & 0 & x & \cdots & x \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x \end{bmatrix}$$

5

其中符号 x 表示通过计算下式表示的矩阵 A_i 的 Cholesky 因式分解得到的非 0 元素：

公式 17

$$A_i = R_i^T R_i$$

10

运动分析器也生成矢量 z_i ，它是通过求解下列公式组得到的：

公式 18

$$R_i^T z_i = d_i$$

15

矩阵 R_i 和矢量 z_i 都是运动分析器的输出参数，这样的参数一起构成适合于在运动系数取消装置 66 中操作的求精运动矢量场的表达式。

20

运动分析器 64 的输出构成了运动系数取消装置 66 的输入。当设置某些求精运动场系数为 0 时取消装置 66 执行的操作包括例如取消那些与能用 \bar{z}' 从 R_i 中去掉的系数对应的元素。所得结果是一个修正了的矩阵 R 和矢量 z 。

可以使用各种方式指定或缺省地隐含段或相邻段组，其中预测运动场是从其导出的。而且，可以使用不同方式来产生预测运动场

00814037.5

说明书 第23/30页

$\Delta x_{\text{prd}}(x, y)$, $\Delta y_{\text{prd}}(x, y)$, 以线性化上面的公式 (11), 以及求解公式组 (18)。

图 4 示出单个视频帧 84, 图中示出被分为多个段 86, 在这里是 30 个。这里的各个段 86 是由 16×16 像素块构成的。并且, 各个段还可以被分割以形成更小的段。这里有些段 86 被分割而形成 8×8 像素块 88。段 86 通常被称为宏块。帧的编码是通过从左到右和从上到下、一个宏块到一个宏块地扫描来执行的。

如上所述, 给定段的运动矢量场服从公式 (8) 中给出的加法运动模型。下面将描述获得预测、求精和最终运动预测场的方式。在示例性实施例中, 运动预测或运动求精场可以是 0。所以, 关于运动矢量场, 可以用各种方式对给定段 S_i 进行编码。例如, 可以仅利用从相邻段外插的预测运动矢量场对段进行编码。或者, 可以利用从相邻段外插的预测运动矢量场及压缩求精运动矢量场一起对段进行编码。或者, 另一方面, 可以仅仅利用压缩运动矢量场对段进行编码, 而不利用预测场。但是, 如果预测场被设为 0, 则发送求精信息。也可以利用零运动矢量场, 例如利用参考帧 $R_{\text{ref}}(x, y)$ 的复制对段进行编码。以及, 例如可以利用没有采用运动矢量场的内部编码 (intra coding) 方法对段进行编码。

在示例性实施例中, 不依赖预测运动矢量场或求精运动矢量场的存在, 给定运动补偿段 S_i 的最终运动矢量场具有由下列公式给出的仿射模型, 其中的上标 i 表示系数与相应段 S_i 相关联的事实。

公式 19

$$\begin{aligned}\Delta x(x, y) &= \beta_1^i + \beta_2^i \cdot (y - y_0^i) + \beta_3^i (x - x_0^i) \\ \Delta y(x, y) &= \beta_4^i + \beta_5^i \cdot (y - y_0^i) + \beta_6^i (x - x_0^i)\end{aligned}$$

其中 x_0^i 和 y_0^i 是段中上面最左边像素的坐标, 而 $\beta_1^i, \dots, \beta_6^i$ 是按照下面所述计算的仿射系数。

00814037.5

说明书 第24/30页

在解码器 14 的示例性实施例中，通过整数精度执行操作。这是通过利用与固定精度相对应的定点实现来取得的。结果，在下文提到的所有系数都是整数值，包括公式 (19) 的系数。在其它实施例中，可以使用其它精度。

- 5 在示例性实施例中，向解码器 14 发送 1 比特，以使用信号通知是否使用相邻段的预测场，但只在至少有一个预测相邻候选时这样作。只有在相邻段 S_k 是非零运动矢量场的情况下，相邻段 S_k 才是用作预测段 S_i 的运动矢量场的候选。

- 10 同样，在这个示例性实施例中，只从当前段的左边和正好上方最近的相邻块来执行预测。所以，相邻段的数量最最多为 4，即，上面两个 8×8 象素块，左边两个 8×8 的象素块。在这个实施例中，只要发送到解码器的比特表示使用从相邻段的预测，就在编码器及解码器中计算预测候选的数量和位置。如果具有例如两个、三个或四个预测候选，则向解码器 14 发送一个或两个选择比特，以表示要
- 15 使用候选。选择信息例如由一个预测方向比特及其后跟着的鉴别比特组成，其中预测方向比特可能存在，也可能不存在，鉴别比特也一样。

- 20 图 5 示出一般以 92 表示的表，其中列出了本发明的示例性实现中选择比特的意义和值。标记 X 表示没有或根据上下文逻辑无关。
- 25 方向比特表示当前正被预测的段的上面或左边是否有候选相邻段，鉴别比特指定两个剩余候选中的哪一个必须被用于运动矢量场的预测。那就是说，当选择上面的或左边的段时，具有两种选择可能性，鉴别比特识别该选择。在表中给出的最后四种情况中，根据最适合候选段的位置，鉴别比特可能存在，也可能不存在。例如，如果方向比特在只有一个候选时表示“从左”，那么就不需要鉴别比特。在解码器 14 中，在对方向比特进行解码之后，胜出候选的方向就已知了。

一旦为当前段的预测选择了相邻段，预测运动矢量场就简单地

00814037.5

说明书 第25/30页

是当前段所覆盖象素区中的相邻段的运动矢量场的外插，表示如下：

公式 20

$$\Delta x_{prd}(x, y) = \beta 1^k + \beta 2^k \cdot (y - y_0^k) + \beta 3^k (x - x_0^k)$$

$$\Delta y_{prd}(x, y) = \beta 4^k + \beta 5^k \cdot (y - y_0^k) + \beta 6^k (x - x_0^k)$$

- 5 其中 x_0^k 和 y_0^k 是相邻段 S_k 中上面最左边象素的坐标，而 $\beta 1^k, \dots, \beta 6^k$ 是与段 S_k 的运动场对应的整数值系数。在公式 20 中，上标 k 表示系数与相邻段 S_k 相关联。

对公式 19 和 20 的分析表示：相邻段 S_k 的运动矢量场通过简单地外插到当前段 S_i 中的象素中而变成段 S_i 的预测运动矢量场。

- 10 求精运动矢量场表现为公式 9 给出的仿射正交模型。但是，在优选实施例中，求精系数被转换为一组辅助求精系数。辅助求精系数允许快速计算最终预测运动场。

- 在优选实施例中，对应于正交仿射基本函数组的公式 9 中的求精系数首先被转换为不同的一组辅助系数。这些系数对应于基本函数组 $\{1, (y-y_0), (x-x_0)\}$ ，其中 x_0, y_0 是段中上面最左边象素的坐标。这个转换被执行，以便取得预测和求精运动矢量场的共同基本函数表达式，即，以便使用相同的基本函数组。通过这样做，根据两组系数的和来计算最终运动矢量场，稍后将对此进行描述。根据求精系数 $c1, \dots, c6$ 来计算段 S_i 的以下辅助系数 $a1, \dots, a6$ 。对于 16×16 象素块的段，计算如下：

公式 21

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 4096 & 6664 & 6664 \\ 0 & -889 & 0 \\ 0 & 0 & -889 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{ 以及 } \begin{bmatrix} a_4 \\ a_5 \\ a_6 \end{bmatrix} = \begin{bmatrix} 4096 & 6664 & 6664 \\ 0 & -889 & 0 \\ 0 & 0 & -889 \end{bmatrix} \cdot \begin{bmatrix} c_4 \\ c_5 \\ c_6 \end{bmatrix}$$

对于 8×8 象素块的段 S_i ，计算形式如下：

00814037.5

说明书 第26/30页

公式 22

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 8192 & 12513 & 12513 \\ 0 & -3575 & 0 \\ 0 & 0 & -3575 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \text{ 以及 } \begin{bmatrix} a_4 \\ a_5 \\ a_6 \end{bmatrix} = \begin{bmatrix} 8192 & 12513 & 12513 \\ 0 & -3575 & 0 \\ 0 & 0 & -3575 \end{bmatrix}$$

结果, 下列整数值位移表示段 S_i 的求精运动矢量场:

5 公式 23

$$\begin{aligned} \Delta x_{refine}(x, y) &= a_1 + a_2 \cdot (y - y_0^i) + a_3(x - x_0^i) \\ \Delta x_{refine}(x, y) &= a_4 + a_5 \cdot (y - y_0^i) + a_6(x - x_0^i) \end{aligned}$$

其中 x_0^i 和 y_0^i 是段 S_i 中上面最左边像素的坐标, 上标 i 表示这些坐标与当前段 S_i 相关联。

10 在示例性实现中, 给定段的最终仿射系数组用以下公式进行计算, 其中所述给定段使用相邻段 S_k 用于运动场预测, 公式中的上标 i 和 k 表示相应系数分别与 S_i 和 S_k 相关联:

公式 24

$$\begin{aligned} \beta_1^i &= a_1 + \Delta x_{pred}(x_0^i, y_0^i) & \beta_4^i &= a_4 + \Delta y_{pred}(x_0^i, y_0^i) \\ \beta_2^i &= a_2 + \beta_2^k & \text{以及} & \beta_5^i &= a_5 + \beta_5^k \\ \beta_3^i &= a_3 + \beta_3^k & & \beta_6^i &= a_6 + \beta_6^k \end{aligned}$$

15 根据整数值系数 β_1, \dots, β_6 , 利用公式 19 产生段 S_i 的最终运动矢量组。下面描述运动矢量被用于从参考帧计算像素强度 (pixel intensities) 的方式。

在示例性实现中, 无论是否期待求精或非求精运动系数, 比特流中运动系数的出现是通过一个比特发信号通知的, 这个比特被称为运动系数指示符 (MCI)

25 同样, 在示例性实现中, 当段 S_i 的运动系数被发送时, 被称为

00814037.5

说明书 第27/30页

运动系数模式 (MCP) 的变长码首先被发送, 以指示哪些系数具有非零值。全零模式是唯一的无效模式, 因为这种可能性可以单独通过 MCI 比特发信号。可以由 MCP 码字表示的有效模式的总数为 63。这是仿射模型的特性。由于它具有 6 个系数, 因此有 2^6 , 即 64 个可能结果。由于零是无效的, 所以 MCP 码字有 63 个可能值。跟在 MCP 码字后面的是由 MCP 模式表示的各个非零运动系数的编码值。各个非零系数的编码值跟在 MCP 码字后面。运动系数 c_j 被编码为幅度变长码字, 表示 c_j 绝对值, 后面跟有表示 c_j 符号的符号比特。在示例性实现中, 使用相同的变长编码表来对不同系数的幅度进行编码。可以使用不同的编码表。零幅度不包括在有效选项中, 因为这种可能性可以由 MCP 码字表示。

通过利用公式 19 计算的最终运动矢量场分量对应于如下面公式所示的离散化步骤:

公式 25

$$D = \frac{1}{65536} = 0.0000152587 \ 890625$$

如果 $(\Delta x(x, y), \Delta y(x, y))$ 表示段 S_i 的最终运动补偿位移, 则在之前帧中的相应非整数坐标为:

公式 26

$$\begin{aligned} x' &= x + \Delta x(x, y) \cdot D \\ y' &= y + \Delta y(x, y) \cdot D \end{aligned}$$

在优选实施例中, 参考帧 R_{ref} 是 $M \times N$ 像素的大小, 具有 $\{0, 1, \dots, 255\}$ 范围的强度值。有效像素坐标 $\{x', y'\}$ 只在范围 $\{0, 1, \dots, M-1\} \times \{0, 1, \dots, N-1\}$ 中被定义。当运动补偿预测需要计算参考帧 R_{ref} 中非整数位置的亮度和色度值时, 采用离散型式的三次卷积内插 (cubic convolution interpolation)。在示例性实现中, 如下所述, 当计算参

00814037.5

说明书 第28/30页

考帧中的重建值时采用定点精度。

首先，对应于段 S_i 中象素 (x, y) 的整数值位移 $(\Delta x(x, y), \Delta y(x, y))$ 用模 65536 的形式表示如下：

公式 27

5

$$\begin{aligned}\Delta x(x, y) &= dx \cdot 65536 + \delta x, \delta x \in \{0, 1, \dots, 65535\} \\ \Delta y(x, y) &= dy \cdot 65536 + \delta y, \delta y \in \{0, 1, \dots, 65535\}\end{aligned}$$

其中 $dx, dy, \delta x$ 和 δy 都是整数值，而后面两个总是非负数。

4×4 的三次卷积窗的 x'_j, y'_k 整数值坐标被定义如下：

公式 28

10

$$\begin{aligned}x'_j &= \text{sat}(x + dx + j - 2, M - 1), j = 1, 2, 3, 4 \\ x'_k &= \text{sat}(y + dy + k - 2, N - 1), k = 1, 2, 3, 4\end{aligned}$$

其中 $\text{sat}(u, v)$ 是如下式表示的饱和函数：

公式 29

15

$$\text{sat}(u, v) = \begin{cases} 0 & u < 0 \\ u & 0 \leq u \leq v \\ v & u > v \end{cases}$$

因此，用于三次卷积的 16 个整数象素值 r_{jk} 如下：

20

公式 29

$$r_{jk} = R_{ref}(x'_j, y'_k) \quad j, k = 1, 2, 3, 4$$

25

其中 x'_j, y'_k 是公式 28 中计算的整数值坐标。

接着，计算卷积系数。在下面，通过舍位的整数除法用 “/” 表

00814037.5

说明书 第29/30页

示, 两者或其操作数总是非负整数。通过利用整数舍位, 下面 u_j 、 v_k ($j, k=1, 2, 3, 4$) 被计算如下:

公式 31

$$\begin{aligned}
 5 \quad u_1 &= spl(\delta x / 256 + 256) & v_1 &= spl(\delta y / 256 + 256) \\
 u_2 &= spl(\delta x / 256) & v_2 &= spl(\delta y / 256) \\
 u_3 &= spl(256 - (\delta x / 256)) \quad \text{以及} & v_3 &= spl(256 - (\delta y / 256)) \\
 u_4 &= 16384 - (u_1 + u_2 + u_3) & v_4 &= 16384 - (v_1 + v_2 + v_3)
 \end{aligned}$$

10 其中 δx 和 δy 是公式 27 的整数值, 而 $spl(s)$ 是正整数变元的整数值函数。

公式 32

$$15 \quad spl(s) = \begin{cases} 16384 - (s^2 \cdot (1280 - 3 \cdot s) + 1024) / 2048 & S \in \{0, 1, \dots, 255\} \\ -((t \cdot (65536 + t^2 - 512 \cdot t) + 1024) / 204 & S \in \{256, \dots, 511\}, t = s - 256 \\ 0 & \text{其它情况} \end{cases}$$

接着计算参考象素值。通过利用舍位的整数除法, 参考象素值被计算如下:

公式 33

$$r = sat \left(\left(\begin{bmatrix} r_{11} & r_{21} & r_{31} & r_{41} \\ r_{12} & r_{22} & r_{32} & r_{42} \\ r_{13} & r_{23} & r_{33} & r_{43} \\ r_{14} & r_{24} & r_{34} & r_{44} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} / 256 \right)^T \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} + 2^{19} \right) / 2^{20}, 255)$$

20

其中公式 30 给出整数值系数 r_{jk} , 公式 31 给出整数值系数 u_j 、 v_k ($j, k=1, 2, 3, 4$), 而公式 29 给出函数 $sat(\dots)$ 。

在运动分析器, 通过采用 $R_{int}(x, y)$ 的一阶泰勒展开来执行线性化步骤:

25 公式 34

00814037.5

说明书 第30/30页

$$\begin{aligned}x'_p &= x_p + \Delta x(x_p, y_p) \\y'_p &= y_p + \Delta y(x_p, y_p)\end{aligned}$$

关于 x 和 y :

公式 35

5

$$\begin{aligned}&R_{ref}(x_p + \Delta x_{prd}(x_p, y_p) + \Delta x_{refine}(x_p, y_p), y_p + \Delta y_{prd}(x_p, y_p) + \Delta y_{refine}(x_p, y_p)) \\&\approx R_{ref}(x'_p, y'_p) + (\Delta x_{refine}(x_p, y_p) + \Delta x_{prd}(x_p, y_p) - \Delta x(x_p, y_p)) \cdot G_x(x'_p, y'_p) \\&\quad + (\Delta y_{refine}(x_p, y_p) + \Delta y_{prd}(x_p, y_p) - \Delta y(x_p, y_p)) \cdot G_y(x'_p, y'_p)\end{aligned}$$

$G_x(x'_p, y'_p)$ 和 $G_y(x'_p, y'_p)$ 是参考帧相对于 x 和 y 的导数值。利用这种近似法，公式 14 中的矩阵 E_i 和矢量 w_i 的元素为：

公式 36

10

$$e_{pk} = \begin{cases} f_k(x_p, y_p) G_x(x'_p, y'_p), & k = 1, 2, \dots, N \\ f_k(x_p, y_p) G_y(x'_p, y'_p), & k = N + 1, N + 2, \dots, N + M \end{cases}$$

公式 37

$$\begin{aligned}w_p &= \ln(x_p, y_p) - R_{ref}(x'_p, y'_p) + G_x(x'_p, y'_p) \Delta x(x_p, y_p) + G_y(x'_p, y'_p) \Delta y(x_p, y_p) - \\&G_x(x'_p, y'_p) \Delta x_{prd}(x_p, y_p) - G_y(x'_p, y'_p) \Delta y_{prd}(x_p, y_p)\end{aligned}$$

15

以上描述针对的是实现本发明的优选示例，本发明的范围不应该限于本描述。本发明的范围由后附权利要求书来定义。

00814037.5

说明书附图

第1/3页

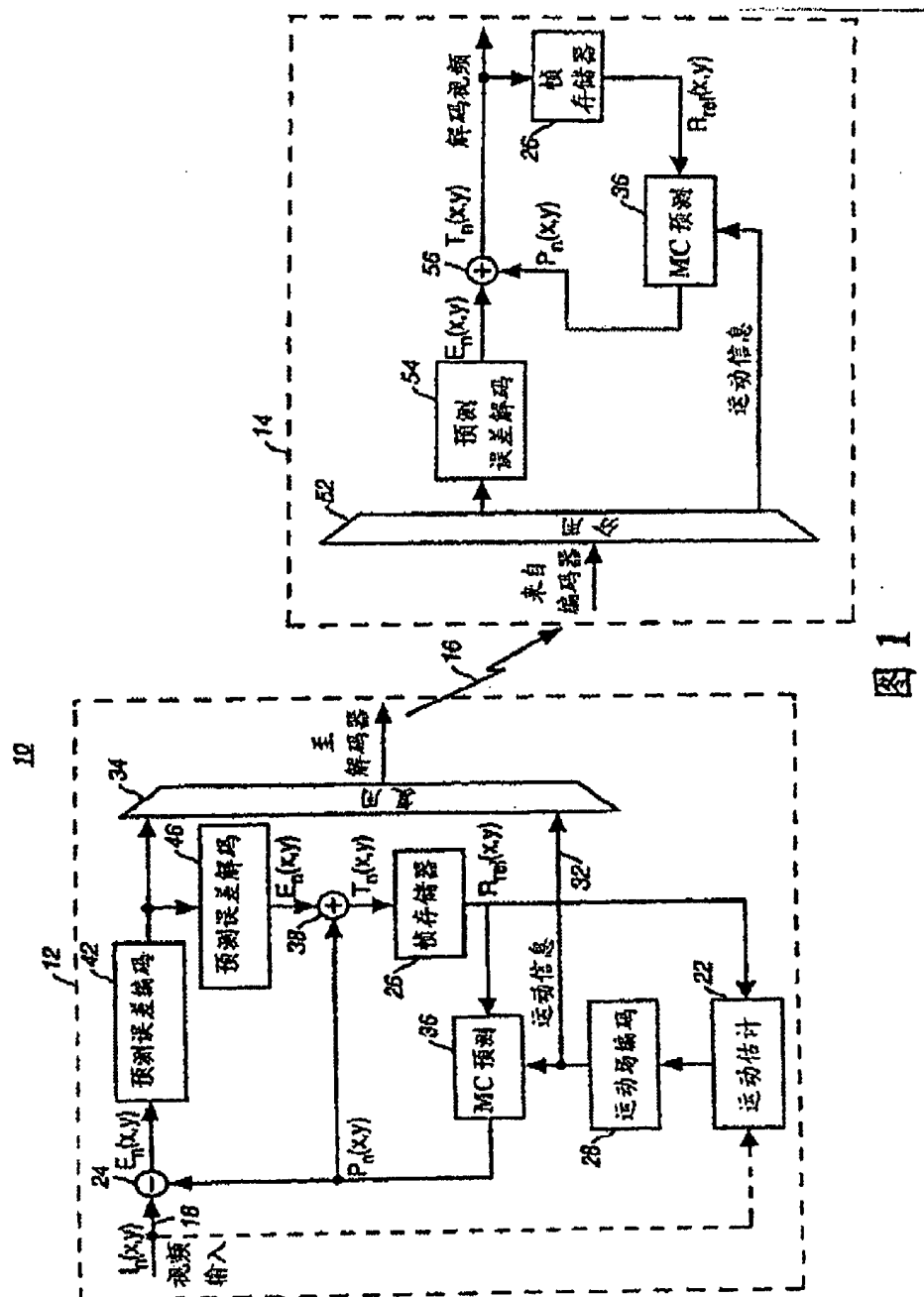
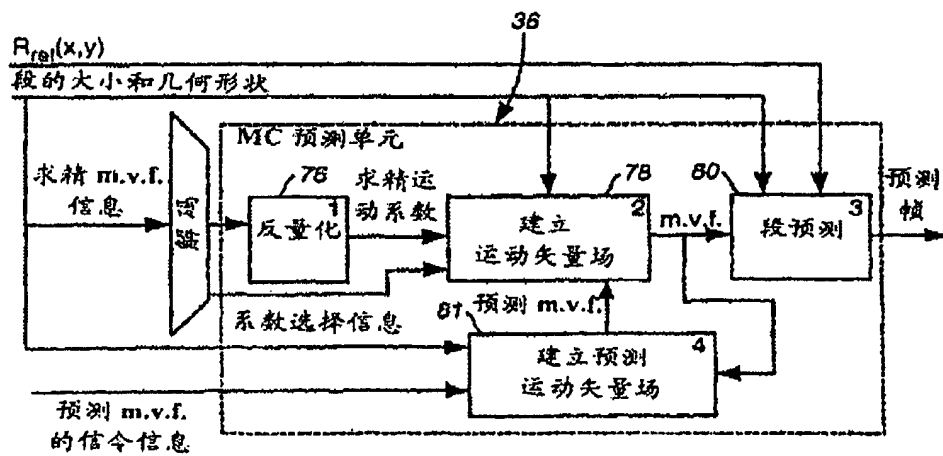
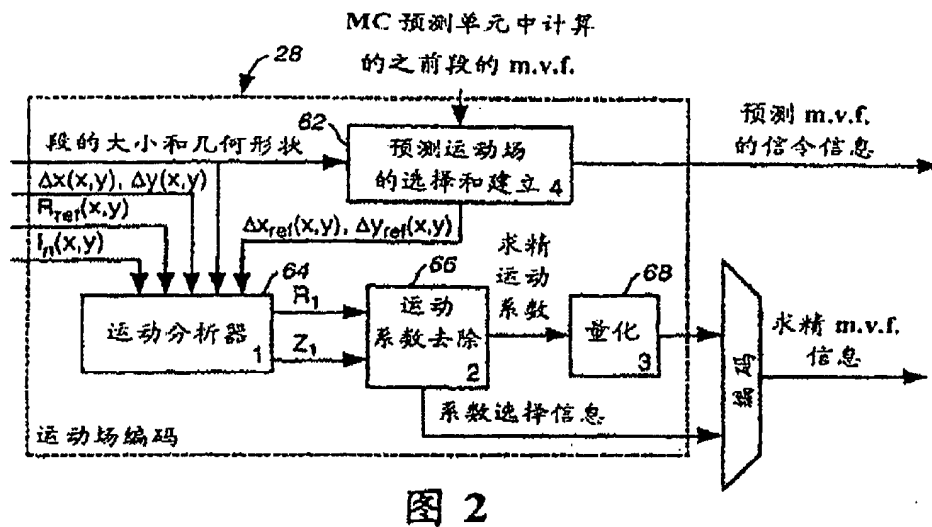


图 1

00814037.5

说明书附图 第2/3页



00814037.5

说明书附图 第3/3页

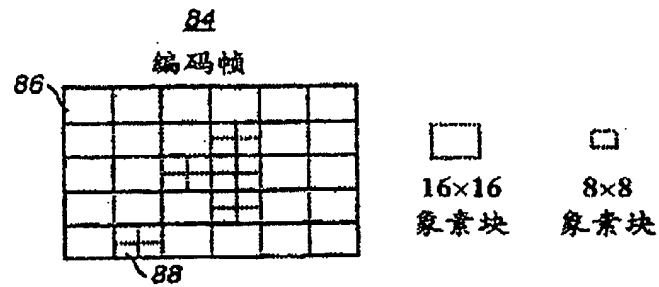


图 4

22

表 1

从上的 候选数	从左的 候选数	出现 方向比特	出现 鉴别比特	胜出候选 的方向	选择 比特数
0	1	X	X	X	无
1	0	X	X	X	无
0	2	X	是	X	1
2	0	X	是	X	1
1	1	是	X	X	1
2	2	是	是	X	2
2	1	是	是	从上	2
2	1	是	X	从左	1
1	2	是	是	从左	2
1	2	是	X	从上	1

图 5